

НАНОЭЛЕКТРОНИКА И НЕЙРОМОРФНЫЕ ВЫЧИСЛИТЕЛЬНЫЕ СИСТЕМЫ

УДК 537.9+004.8

КОМПАКТНАЯ МОДЕЛЬ ОПИСАНИЯ ПЛАСТИЧНОСТИ МЕМРИСТОРОВ НА ОСНОВЕ НАНОСЛОЕВ LiNbO_3 И КОМПОЗИТА $(\text{Co-Fe-V})_x(\text{LiNbO}_3)_{100-x}$ ПО БИОПОДОБНОМУ ПРАВИЛУ STDP

© 2024 г. А. В. Емельянов^{1,2,*}, К. Э. Никируй¹, А. И. Ильясов^{1,3}, Е. А. Цывкунова⁴,
И. А. Суражевский¹, В. А. Демин¹, Я. Линь⁵, Е. Тао⁵, В. В. Рыльков^{1,6}

¹Национальный исследовательский центр “Курчатовский институт”, Москва, Россия

²Московский физико-технический институт, Долгопрудный, Россия

³Московский государственный университет им. М.В. Ломоносова, Москва, Россия

⁴Национальный исследовательский ядерный университет “МИФИ”, Москва, Россия

⁵Северо-восточный педагогический университет, Чанчунь, Китай

⁶Институт теоретической и прикладной электродинамики РАН, Москва, Россия

*E-mail: emelyanov.andrey@mail.ru

Поступила в редакцию 28.11.2023 г.

После доработки 22.12.2023 г.

Принята к публикации 22.12.2023 г.

Предложена компактная феноменологическая модель описания пластичности мемристорных структур на базе нанослоев LiNbO_3 и композита $(\text{Co-Fe-V})_x(\text{LiNbO}_3)_{100-x}$, учитывающая особенности резистивного переключения структур и прыжкового электронного транспорта в аморфном LiNbO_3 . Модель хорошо описывает вольт-амперные характеристики (ВАХ) мемристорных кроссбар-массива, а найденные при подгонке ВАХ эффективные микроскопические параметры позволяют предсказать результат изменения проводимости согласно пластичности, зависящей от времени прихода парных импульсов (spike-timing-dependent plasticity – STDP), а кроме того, зависимость окна STDP от начальной проводимости мемристора. Полученные результаты могут быть использованы при разработке алгоритмов обучения спайковых нейроморфных вычислительных систем и выявления мемристорного STDP, эффективного для их реализации.

DOI: 10.56304/S1992722323601027

ВВЕДЕНИЕ

Актуальность исследования мемристорных обусловлена, в первую очередь, возможностью их применения в нейроморфных вычислительных системах (НВС) [1, 2]. Данные системы, состоящие из искусственных нейронов и синаптических связей между ними на основе массивов кроссбар-мемристорных, могут обеспечить высокую производительность и существенно большую энергоэффективность при решении задач искусственного интеллекта, чем традиционные компьютеры с архитектурой фон Неймана [3]. Для реализации НВС требуются матрицы кроссбар-мемристорных с весьма большим числом элементов (уже сейчас для демонстрации НВС используются мемристорные матрицы размером 256×256 и более [4, 5]). По этой причине в создании интегральных схем нейроморфной электроники (ИСНЭ) с оптимальной архитектурой все более актуальным становится разработка компактных моделей базовых компонентов ИСНЭ, в частности мемристорных.

Термин “компактная” модель предполагает, в первую очередь, вычислительную простоту, чрезвычайно важную для моделирования одно-временной работы миллионов и миллиардов элементов ИСНЭ по аналогии с подходом, используемым в комплементарной кремниевой микроэлектронике. В данном случае такой подход существенно повышает производительность разработки интегральных схем, обладает большей сходимостью при работе алгоритмов систем автоматизированного проектирования и имеет приемлемый рабочий диапазон параметров в сравнении со сложными физическими моделями транзисторов [6].

Особый интерес в настоящее время проявляется к импульсным (спайковым) НВС, в которых может быть достигнута наибольшая энергоэффективность, приближающая их к биологическим прототипам [7]. В этом случае важен аналоговый или, по крайней мере, многоуровневый характер резистивного переключения (РП) мемристорных устройств. При этом информация

кодируется в частотах [8] или временах [9] прихода импульсов (спайков), а синаптический вес (проводимость мемристора) изменяется согласно биоподобным правилам, в частности когда пластичность синапса зависит от времени прихода парных спайков (spike-timing-dependent plasticity – STDP) [10].

Несмотря на большое количество работ по демонстрации STDP в мемристивных структурах [11–14], не существует универсального подхода для описания зависимости окна мемристивной STDP, которое может существенно отличаться от математических (поведенческих) форм STDP, используемых в программных реализациях импульсных нейронных сетей [15]. Отметим, что для корректного моделирования аппаратной спайковой НВС важно учитывать зависимость STDP от начальной проводимости мемристора. Из общих соображений ясно, что для решения отмеченных задач необходимо развитие моделей, учитывающих физические особенности РП, в частности возможность образования при РП многочисленных филаментов, способствующих увеличению степени пластичности мемристора или многоуровневому характеру его переключения. Многоуровневые РП наблюдаются, в частности, в мемристорах с перколяционной проводимостью на базе нанокompозита парилена–Ag [16] или $(\text{Co-Fe-V})_x(\text{LiNbO}_3)_{100-x}$ [17, 18]. В последних также важную роль играют прыжковый электронный транспорт и эффекты резонансного туннелирования [19] в аморфной прослойке LiNbO_3 , формируемой у нижнего электрода структуры.

В [20] была развита модель РП нанокompозитных (НК) мемристоров, учитывающая как распределение филаментов по значениям проводимостей, так и особенности электронного транспорта в аморфной прослойке, что позволило описать характер поведения их пластичности (многоуровневого РП), в том числе зависимость изменения проводимости ΔG мемристора от начальной его проводимости G_0 при воздействии парных импульсов, т.е. по правилу типа STDP. При этом важно, что по вычислительной сложности (эффективности) модель [20] ввиду использования в ее основе реалистичных физических механизмов уступает более простым поведенческим моделям (например, [21, 22]), в которых функциональные уравнения выводятся не столько из физических соображений, сколько через выбор подходящих функций и подбор параметров для наилучшей аппроксимации поведения реального устройства.

Развитая в [20] модель РП была апробирована на примере описания мемристивных свойств структур металл/нанокompозит/металл (М/НК/М, где НК – $(\text{Co-Fe-V})_x(\text{LiNbO}_3)_{100-x}$, $x \approx 10-15$ ат. %) с самоорганизованной аморфной прослойкой LiNbO_3

(LNO), формируемой у нижнего электрода на начальном этапе роста НК при повышенном парциальном давлении кислорода ($P_{\text{O}_2} \approx 2.5 \times 10^{-5}$ Торр). В этом случае для корректного описания ΔG принципиально важным оказался учет распределения проводимостей (длин) филаментов (или зазоров между границей высокопроводящей части филамента и нижним электродом структуры М/НК/М [17]). С другой стороны, межграницный слой LNO в образцах М/НК/М можно формировать целенаправленно путем распыления мишени из монокристаллического LNO, а затем осадить НК. При структурных исследованиях данных образцов методами высокоразрешающей электронной микроскопии выявлено, что в данном случае четко проявляется граница раздела НК/LNO, шероховатость которой не превышает размера гранул (~2–3 нм). Между тем в случае самоорганизованной прослойки граница НК/LNO оказывается существенно более рельефной, причем размер гранул на границе достигает $a_g \approx 6$ нм (подробный сравнительный анализ структуры данных образцов приведен [23]). Поэтому представляется естественным попытаться описать пластичность мемристоров М/НК/LNO/М со встроенной межграницной прослойкой в рамках модели эффективного филамента или многих филаментов “одинаковой” длины, т.е. с узким их распределением по проводимости [20].

В настоящей работе продемонстрирован компактный феноменологический подход для упрощенного физического описания поведения проводимости по правилам STDP при разных начальных сопротивлениях элементов кроссбар-матрицы на базе структур М/ $(\text{CoFeV})_x(\text{LiNbO}_3)_{100-x}$ /LNO/М ($x \approx 10-15$ ат. %) с использованием параметров подгонки вольт-амперных характеристик (ВАХ) элементов в рамках модели эффективного филамента при конечном сопротивлении шин матрицы. По вычислительной эффективности данная модель оказывается сравнимой с поведенческими моделями, предсказывая поведение окна STDP, что открывает возможность моделирования больших импульсных НВС с целью выбора их оптимальных архитектур для аппаратной реализации.

ОБРАЗЦЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

Исследовали массивы 16×16 мемристоров М/НК/LNO/М, изготовленные в кроссбар-геометрии, где отдельные элементы устройства находятся на пересечении перпендикулярных друг другу шин, выполняющих роль верхних и нижних электродов структуры. Структуры изготавливали с помощью технологий лазерной фотолитографии (литограф Heidelberg 66fs разрешением 1 мкм), плазмо-химического осаждения (Trion Oracle III)

и ионно-лучевого распыления двух мишеней: из кристаллического LiNbO_3 и составной мишени из пластины $\text{Co}_{40}\text{Fe}_{40}\text{V}_{20}$ с неравномерно расположенными навесками LiNbO_3 , что позволяло в едином цикле формировать кроссбар-элементы с различным содержанием металла в диапазоне $x = 10\text{--}25$ ат. % [24]. Толщины слоев LiNbO_3 и НК $(\text{Co-Fe-V})_x(\text{LiNbO}_3)_{100-x}$ в структурах составляли 10 и 230 нм соответственно.

Электродные шины из Au имели значительную толщину (150 и 350 нм для нижней и верхней шин соответственно), в связи с чем могут возникать краевые эффекты, приводящие к токовым утечкам и локальному электрическому пробую структур. Чтобы избежать появления этих эффектов, рабочую область НК изолировали от нижней шины слоем Si_3N_4 толщиной 40 нм с использованием плазмо-химического осаждения при температуре 300°C . Дополнительно НК изолировали от верхней шины слоем Si_3N_4 толщиной 6–7 нм. Размер окон в слоях Si_3N_4 (17×17 мкм²) был выбран заметно меньшим, чем ширина нижней и верхней шин (20 мкм). В этих условиях краевые эффекты практически не проявлялись.

Оптимальное содержание металла в слое НК для наблюдения РП составляло $x \approx 20$ ат. %. При этом суммарное сопротивление шин (около 50 Ом) было сравнимо с сопротивлением структур в низкоомном состоянии (~60 Ом). Поэтому представленные ниже эффективные мемристивные характеристики кроссбар-структур отличаются от тех характеристик, что раньше наблюдали для отдельных конденсаторных мемристоров М/НК/М [17, 18].

Исследования ВАХ и мемристивных свойств структур проводили при комнатной температуре с помощью источника-измерителя NI PXIe-4140 (National Instruments) и аналитической зондовой станции PM5 (Cascade Microtech) с оптической системой микрометрического перемещения зондов PSM-100 (Motic). ВАХ измеряли в импульсном режиме при заземленном нижнем электроде по линейно-ступенчатому закону изменения амплитуды импульсов напряжения смещения U верхнего электрода в последовательности от $0 \rightarrow +U_0 \rightarrow -U_0 \rightarrow 0$ В ($U_0 = 5\text{--}6$ В) с шагом 0.1 В при длительности импульсов 50 мс. Импульсы подавали непрерывно ($\delta t = 0$) или с некоторой задержкой ($\delta t = 50, 100, 500$ мс), что позволяло в широком диапазоне изменять время одного цикла измерения ВАХ (10–110 с) при неизменном количестве точек дискретизации (вставка на рис. 2).

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Численное описание ВАХ проводили на основе развитой в [20] модели многофиламентного

РП, учитывающей ранее выявленные особенности структур М/НК/М [17]. Согласно этой модели, РП НК-структур происходят в тонком высокоомном слое LNO (зазоре, отделяющем низкоомные филаменты от нижнего электрода структуры). В начальном состоянии после изготовления структуры диспергированные атомы металла (Co, Fe) равномерно распределены в слое НК [17]. После подачи напряжения и протекания тока возможна нуклеация диспергированных атомов вокруг перколяционных цепочек из гранул. Когда к верхнему электроду прикладывается достаточно большое положительное напряжение, структура переходит в низкоомное состояние за счет электромиграции вакансий и катионов в сторону нижнего электрода, что приводит к уменьшению эффективного диэлектрического зазора в слое LNO. Обратная ситуация возникает, когда на структуру подается достаточно большое отрицательное напряжение. В этом случае структура переключается в высокоомное состояние из-за перемещения кислородных вакансий и катионов к верхнему электроду вдоль цепочек и увеличения эффективного зазора между филаментами в LNO и нижним электродом.

Проводимость структуры в промежуточном резистивном состоянии носит прыжковый характер, определяемый большой плотностью локализованных состояний вблизи уровня Ферми. В этих условиях могут возникать эффекты резонансного туннелирования [18, 19, 25], которые вызывают экспоненциальную зависимость удельного сопротивления зазора от его толщины:

$$r_g \sim \exp \frac{2b}{a(n+1)}, \quad (1)$$

где b – толщина зазора, a – радиус локализованного состояния, а n – количество резонансных энергетических состояний в цепочке локализованных дефектов.

При расчетах ВАХ полагали, что формируемые в LNO филаменты обладают узким распределением по их проводимости [20], т.е. пренебрегли распределением филаментов по длинам и подвижностям в них вакансий/ионов металла. Однако, когда свели описание сложной системы к модели “эффективного филамента”, получили весьма удовлетворительное описание экспериментальных данных при сохранении вполне физического характера модели.

В модели эффективного филамента ток вычисляли по закону Ома, исходя из напряжения на структуре М/НК/LNO/М и ее сопротивления, которое складывается из сопротивлений НК-слоя и диэлектрического зазора:

$$R = \left(\rho_f (d - b) + \rho_g b \exp \frac{2b}{a_{\text{eff}}} \right) / S_f n_f,$$

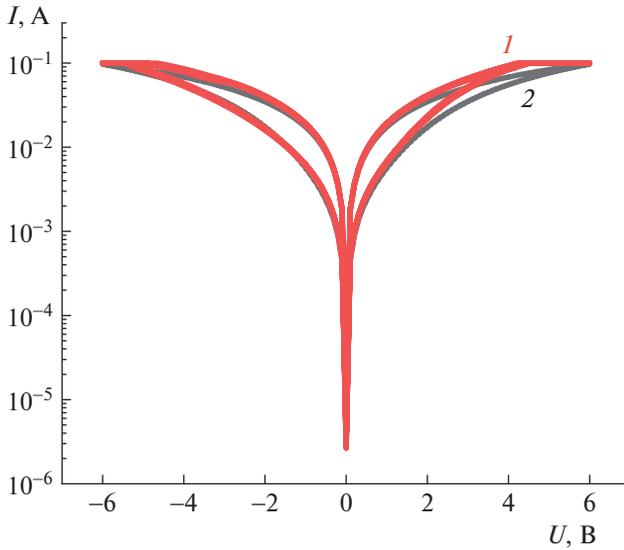


Рис. 1. Экспериментальные (1) ВАХ мемристивных структур и численное моделирование (2).

$$a_{\text{eff}} = a(n+1),$$

где ρ_f , ρ_g – коэффициенты размерности удельного сопротивления, определяющие падения напряжения на НК-слое и зазоре, d – толщина активной области (НК и LNO), S_f – площадь сечения филамента, n_f – количество филаментов.

При расчете поля в зазоре учитывали перераспределение напряжения между НК-слоем и зазором, последнее выражается в виде

$$V_g = V_1 / (1 + \rho_f(d-b) / \rho_g b e^x),$$

$$x = 2b/a_{\text{eff}},$$

где V_1 – напряжение с учетом тока ограничения, заданного в эксперименте (дополнительно напряжение перераспределяется за счет конечного сопротивления шин к кроссбар-структурам). При этом величина эффективного зазора изменяется за счет дрейфа вакансий в поле [19]:

$$\Delta b \approx -\mu(V_g/b)\Delta t,$$

где μ – подвижность вакансий или нижней границы филамента, Δt – шаг моделирования.

На рис. 1 представлены экспериментальные кривые ВАХ для мемристивных кроссбар-структур и результаты описания ВАХ в рамках модели

Таблица 1. Значение параметров, использованных при моделировании

a , нм	n	d , нм	Δt , мс	n_f	S_f , нм ²	ρ_g , Ом м	ρ_f , Ом м
1	6	180	0.1	10^8	25	0.02	0.43

эффективного филамента для параметров, приведенных в табл. 1, при которых достигается точность подгонки 4% (для небольших напряжений порядка напряжений чтения). Иными словами, предложенный подход вполне удовлетворительно описывает экспериментальные данные. Отметим, что при построении компактных аналитических моделей процессов РП обычно учитывают локальный джоулев разогрев и вводят дополнительно уравнение теплопроводности для описания распространения тепла и нагрева активной области мемристора ([20] и ссылки в ней). По-видимому, наблюдаемое расхождение экспериментальных и расчетных кривых ВАХ (при относительно высоких напряжениях смещения $\sim 2-4$ В) может быть обусловлено наличием не учитываемого в настоящей работе джоулева разогрева высокопроводящих филаментов, ввиду которого подвижность составляющих филамент вакансий кислорода и ионов металла может изменяться экспоненциально. При подгонке использовали эффективное значение $\mu = 4.78 \times 10^{-5} \text{ см}^2/(\text{В} \cdot \text{с})$.

Значение подвижности в изучаемой модели выбрано для филамента с некоторой средней или эффективной температурой, причем очевидно, что разогрев увеличивается с ростом прикладываемого напряжения. По этой причине при высоких напряжениях соответствующие изменения проводимости мемристора в эксперименте оказываются больше, чем при численном моделировании (рис. 1), однако расхождение не очень велико. Что касается влияния нестационарных температурных процессов, то они в данном случае не проявляются при длительностях пилообразной развертки 10–110 с, которые существенно превосходят характерные времена установления термодинамического равновесия. Действительно, из рис. 2 видно, что кривые ВАХ при различных задержках между импульсами пилообразной развертки напряжения практически неразличимы.

Основой для реализации импульсных нейронных сетей являются биоподобные правила изменения синаптического веса типа STDP. В эксперименте изменение проводимости осуществлялось за счет приложения парных импульсов к нижнему и верхнему электродам структуры, имитирующих спайки от пре- и постнейронов с различным межспайковым интервалом. Пример экспериментально измеренного окна STDP, т.е. зависимости изменения проводимости мемристора от межспайкового интервала показан на рис. 3. Также было проведено моделирование окна STDP для параметров, найденных по ВАХ. Из представленных данных следует, что моделирование окна STDP для начального состояния мемристора $G_0 = 7 \text{ мСм}$ ($R_0 \approx 140 \text{ Ом}$) позволяет с хорошей точностью прогнозировать изменение веса

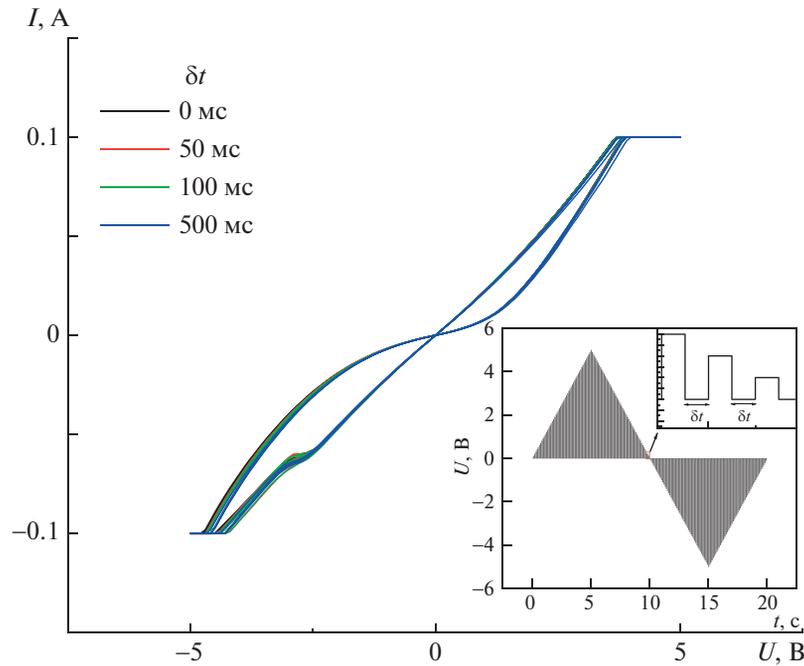


Рис. 2. Экспериментальные ВАХ с различной скважностью измерений. На вставке показана развертка по напряжению, использованная при измерении ВАХ.

(среднеквадратичное отклонение составляет 2.9%).

Для моделирования нейроморфных систем с мемристивными синапсами необходима корректная модель изменения синаптического веса

не только относительно межспайкового интервала, но и от начального состояния мемристора, так как мемристивная STDP демонстрирует довольно сильную зависимость от его начального состояния, отличную от мультипликативного характера STDP [17]. Экспериментальная и расчетная зависимости максимальной амплитуды окна STDP от начального состояния мемристора показаны на рис. 4. Видно, что амплитуда окна STDP может уменьшаться при изменении G_0 более чем на порядок, что существенно при моделировании работы импульсных НВС.

ЗАКЛЮЧЕНИЕ

Модель эффективного филамента хорошо описывает вольт-амперные характеристики матричных кроссбар-мемристоров на основе слоистых структур нанокompозит–аморфный оксид, а найденные при подгонке ВАХ эффективные микроскопические параметры мемристора позволяют предсказать результат изменения проводимости согласно биоподобному правилу типа STDP, а кроме того, зависимость окна STDP от начальной проводимости мемристора. Полученные результаты могут быть использованы при разработке алгоритмов обучения масштабируемых спайковых нейроморфных вычислительных систем и выявления мемристивного STDP, эффективного для их реализации.

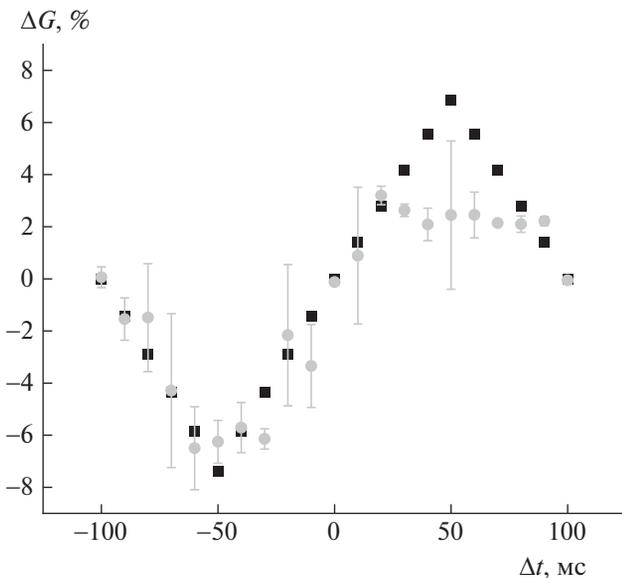


Рис. 3. Окно STDP: экспериментальные данные (серые точки) и результат моделирования (черные квадраты) при начальном состоянии $G_0 = 7$ мСм.

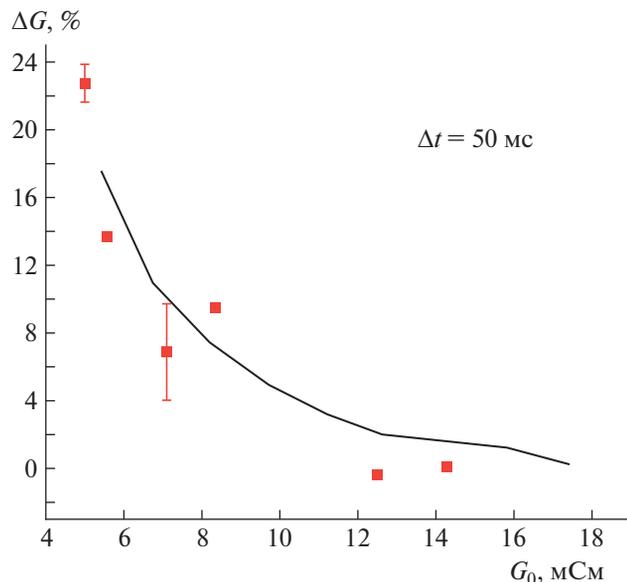


Рис. 4. Изменение веса согласно правилу STDP: результат моделирования (черная кривая) и экспериментальные данные (красные квадраты) при различных начальных состояниях G_0 .

Работа выполнена при финансовой поддержке Российского научного фонда (проект № 22-19-00171) в части создания нанокompозитных структур и моделирования их мемристивных свойств. Экспериментальные исследования процессов резистивного переключения структур выполнены при поддержке государственного задания НИЦ “Курчатовский институт” (№ 86) с использованием оборудования Ресурсных центров.

СПИСОК ЛИТЕРАТУРЫ

- Zhang Y., Wang Z., Zhu J. et al. // Appl. Phys. Rev. 2020. V. 7. № 1. P. 011308. <https://doi.org/10.1063/1.5124027>
- Zhou G., Wang Zh., Sun B. et al. // Adv. Electron. Mater. 2022. V. 8. № 7. P. 2101127. <https://doi.org/10.1002/aelm.202101127>
- Wang Z., Wu H., Burr G.W. et al. // Nat. Rev. Mater. 2020. V. 5. № 3. P. 173. <https://doi.org/10.1038/s41578-019-0159-3>
- Rao M., Tang H., Wu J. et al. // Nature. 2023. V. 615. P. 823. <https://doi.org/10.1038/s41586-023-05759-5>
- Zhang W., Yao P., Gao B. et al. // Science. 2023. V. 381. P. 1205. <https://doi.org/10.1126/science.ade3483>
- Gasseling T. // Microwave J. 2012. V. 55. P. 74.
- Dora S., Kasabov N. // Big Data Cogn. Comput. 2021. V. 5. P. 67. <https://doi.org/10.3390/bdcc5040067>
- Prezioso M., Mahmoodi M.R., Bayat F.M. et al. // Nat. Commun. 2018. V. 9. № 1. P. 5311. <https://doi.org/10.1038/s41467-018-07757-y>
- Никуруй К.Э., Емельянов А.В., Ситников А.В. и др. // Российские нанотехнологии. 2021. Т. 16. № 6. С. 761.
- Caporale N., Dan Y. // Annu. Rev. Neurosci. 2008. V. 31. P. 25. <https://doi.org/10.1146/annurev.neuro.31.060407.125639>
- Prezioso M., Merrikkh Bayat F., Hoskins B. et al. // Sci. Rep. 2016. V. 6. P. 21331. <https://doi.org/10.1038/srep21331>
- Matsukatova A.N., Prudnikov N.V., Kulagin V.A. et al. // Adv. Intell. Syst. 2023. V. 5. № 6. P. 2200407. <https://doi.org/10.1002/aisy.202200407>
- Wang Z., Zeng T., Ren Y. et al. // Nat. Commun. 2020. V. 11. № 1. P. 1510. <https://doi.org/10.1038/s41467-020-15158-3>
- Koryazhkina M.N., Ryabova M.A., Okulich E. V. et al. // Phys. Status Solidi. A. 2023. V. 2. P. 2200742. <https://doi.org/10.1002/pssa.202200742>
- Sboev A., Vlasov D., Serenko A. et al. // AIP Adv. 2016. V. 6. № 11. P. 111305. <https://doi.org/10.1063/1.4967353>
- Matsukatova A.N., Vdovichenko A.Yu., Patsaev T.D. et al. // Nano Res. 2023. V. 16. P. 3207. <https://doi.org/10.1007/s12274-022-5027-6>
- Martyshov M.N., Emelyanov A.V., Demin V.A. et al. // Phys. Rev. Appl. 2020. V. 14. № 3. P. 034016. <https://doi.org/10.1103/PhysRevApplied.14.034016>
- Demin V.A., Nekhaev D.V., Surazhevsky I.A. et al. // Neural Networks. 2021. V. 134. P. 64. <https://doi.org/10.1016/j.neunet.2020.11.005>
- Glazman L.I., Matveev K.A. // Sov. Phys. JETP. 1988. V. 67. № 6. P. 1276.
- Дёмин В.А., Ильясов А.И., Рьльков В.В. и др. // Российские нанотехнологии. 2023. Т. 18. № 2. С. 276.
- Kvatinsky S., Ramadan M., Friedman E.G., Kolodny A. // IEEE Trans. Circuits Syst. II: Express Briefs. 2015. V. 62. № 8. P. 786. <https://doi.org/10.1109/TCSII.2015.2433536>
- Ryndin E., Andreeva N., Luchinin V. // Micromachines. 2022. V. 13. № 1. P. 98. <https://doi.org/10.3390/mi13010098>
- Никуруй К.Э. // Дисс. ... канд. физ.-мат. наук, М.: НИЦ “Курчатовский институт”, 2022. 136 с.
- Matsukatova A.N., Pyasov A.I., Nikiruy K.E. et al. // Nanomaterials. 2022. V. 12. № 19. P. 3455. <https://doi.org/10.3390/nano12193455>
- Lutsev L.V., Kalinin Y.E., Sitnikov A.V., Stognei O.V. // Phys. Solid State. 2002. V. 44. № 10. P. 1889. <https://doi.org/10.1134/1.1514778>